Devillers, R., Y. Bedard & M. Gervais, 2004, Indicateurs de qualité pour réduire les risques de mauvaise utilisation des données géospatiales, Revue Internationale de Géomatique,Vol. 14, No. 1, p. 35-57

Indicateurs de qualité pour réduire les risques de mauvaise utilisation des données géospatiales

Rodolphe Devillers *, — Yvan Bédard* — Marc Gervais *,****

** Centre de recherche en géomatique (CRG), Université Laval*
Pavillon Casault, Québec G1K 7P4 (QC), Canada
rodolphe.devillers.1@ulaval.ca ; [yvan.bedard],[marc.gervais]@scg.ulaval.ca

*** Institut Francilien des GéoSciences, Université de Marne-la-Vallée*
5 Boulevard Descartes, 77454 Marne la Vallée, France

RÉSUMÉ. *Les utilisateurs de données géospatiales doivent être conscients de la qualité des données qu'ils manipulent afin de réduire les risques de mauvaises utilisations. L'information décrivant la qualité est variée et peut être représentée à différents niveaux de détail. Les utilisateurs peuvent donc accéder à de grands volumes d'information sur la qualité et se retrouver perdus dans cette abondance d'information. Cet article propose l'utilisation d'indicateurs de qualité pour améliorer la communication des informations relatives à la qualité des données géospatiales. Les concepts de tableau de bord et d'indicateur sont présentés et adaptés au domaine géospatial pour être intégrés dans des SIG. Un aperçu d'un prototype nommé Manuel à l'Usager Multidimensionnel (MUM) communiquant des indicateurs de qualité dans une interface de type SIG est présenté.*

ABSTRACT. *Geospatial data users have to be aware of the data quality in order to reduce the risks of data misuse. Information describing quality is diverse and can be represented at different levels of detail. Users can then access huge volumes of quality information, being lost in this information overload. This paper suggests the use of quality indicators to enhance the communication of geospatial data quality information. Concepts of dashboard and indicator are presented and adapted to the geospatial domain to be integrated into a GIS interface. An overview of a prototype named Multidimensional User Manual (MUM) that communicates quality indicators into a GIS interface is presented.*

MOTS-CLÉS : *SIG, qualité des données, indicateurs, tableau de bord, risques, OLAP.*

KEYWORDS : *GIS, data quality, indicators, dashboard, risks, OLAP.*

1. Introduction

Les domaines utilisant des données géospatiales sont variés et de nouvelles applications émergent fréquemment (Longley *et al.*, 1999). Si l'utilisation de données géospatiales était il y a quelques années le domaine d'usagers experts utilisant des systèmes complexes et onéreux, la réalité a changé de manière significative (Hunter, 1999). Les logiciels SIG et les données géospatiales sont désormais accessibles à de faibles coûts, voire gratuitement sur Internet et sont de plus en plus faciles d'utilisation (Goodchild, 1995; Agumya et Hunter, 1997; Curry, 1998; Elshaw Thrall et Thrall, 1999). Les données géospatiales ne sont plus uniquement manipulées au niveau opérationnel dans les organisations mais également aux niveaux stratégique et tactique (Longley *et al*., 1999). Elles sont maintenant de plus en plus utilisées pour supporter les processus de prise de décision (Hunter, 1999), allant de la sélection d'itinéraires pour planifier ses vacances, à la gestion d'un réseau routier par des agences gouvernementales. Avec le développement de services basés sur la localisation (LBS) et des technologies sans fils, il sera probablement habituel pour tout le monde dans un proche futur de prendre des décisions basées sur des données géospatiales visualisées sur des téléphones portables, systèmes nomades (ex : Palm Pilot), systèmes de navigation dans les voitures, etc.

Ces changements dans le contexte dans lequel les données géospatiales sont utilisées accroissent significativement les risques de mauvaises utilisations de ces données (Epstein *et al.*, 1998). Ainsi, Goodchild (1995) dit que « les SIG sont leurs propres ennemis : en invitant les gens à trouver de nouvelles utilisations des données, cela les invite aussi à être irresponsables dans leur utilisation» (traduction libre). Des cas de mauvaises utilisations sont fréquemment cités dans les revues scientifiques, les médias et les cas de jurisprudence (Blackmore, 1985; Beard, 1989; Goodchild et Kemp, 1990; Monmonier, 1994; Curry, 1998; Gervais, 2004). Beard (1989) identifie les mauvaises utilisations de données géospatiales comme étant des erreurs d'utilisation (*use error*), les ajoutant aux deux types d'erreur fréquemment mentionnés : les erreurs d'acquisition et les erreurs de traitement (*source errors* et *process errors)*. Cette problématique favorise l'émergence de travaux visant à offrir des SIG pouvant prendre en compte la qualité des données manipulées (*quality-aware* ou *error-aware GIS*) (Buttenfield, 1993; Duckham et McCreadie, 1999; Faïz, 1999; Hunter et Reinke, 2000; Duckham et McCreadie, 2002; Gan et Shi, 2002; Qiu et Hunter, 2002; Reinke et Hunter, 2002). Le problème est approché de différentes manières souvent complémentaires entre elles. Hunter et Reinke (2000) proposent de fournir des avertissements aux utilisateurs de SIG lorsque ceux-ci effectuent des « opérations illogiques ». Cette approche utilisant une base de règles s'adresse aux problèmes résultant de la manipulation de données avec les fonctions des SIG. Cette approche est de notre point de vue nécessaire mais n'apporte qu'une solution partielle au problème. En effet, beaucoup d'utilisateurs emploient les SIG à des fins de visualisation, sans utiliser d'opérateurs d'analyse (Lardon *et al.*, 2001; Roche, 2001). L'utilisateur du SIG peut ainsi effectuer des analyses de façon cognitive,

comme identifier visuellement le chemin à suivre entre deux points de la carte, localiser la borne-fontaine la plus proche d'un bâtiment ou encore compter le nombre de bâtiments bordant une certaine rivière. Dans ces cas, des messages d'avertissement communiqués automatiquement ne pourraient pas prévenir les mauvaises utilisations. D'autre part, la base de règles nécessaire serait difficilement exhaustive, cohérente, et adaptée aux différents profils des utilisateurs, certaines règles pouvant par exemple être contextuelles à l'expertise ou au domaine d'application des utilisateurs. Cette approche ne peut donc pas prendre en compte tous les types de mauvaises utilisations. Différentes solutions peuvent alors être explorées pour réduire les risques de mauvaises utilisations des données géospatiales. Par exemple, Krek et Frank (1999) recommandent la création de jeux de données créés spécifiquement pour certains types d'utilisation (ex *:* navigation pédestre dans des environnements urbains), validant ainsi dès la production du jeu de données l'adéquation de celles-ci à leur utilisation. Si cette approche existe depuis longtemps (ex : les données topographiques étaient initialement produites pour des applications militaires), le contexte a changé. Les cartes de base (ex : topographique, cadastrales) sont à présent souvent utilisées en dehors de leur objectif initial. D'un autre côté, basé sur une exploration des considérations légales reliées à l'utilisation de systèmes d'information géographiques (SIG), Gervais (2004) recommande, entre autres, aux utilisateurs novices le recours à des experts en géomatique (*geomatics officer*) qui identifieraient les risques potentiels de mauvaises utilisations et confirmeraient, ou infirmeraient, l'adéquation de certains jeux de données pour certaines utilisations. Une autre manière pouvant permettre la réduction des risques de mauvaises utilisations, présentée dans cet article, consiste à fournir au sein de l'interface du SIG des informations contextuelles et compréhensibles sur la qualité des jeux de données manipulés. L'utilisateur étant informé de la qualité, si celle-ci semble problématique, il a alors le choix entre rechercher des données répondant mieux à ses besoins (réduisant ainsi l'incertitude) ou utiliser les données en étant conscient des conséquences éventuelles (absorbant ainsi l'incertitude résiduelle). Les métadonnées (*i.e.* données sur les données) distribuées par certains producteurs de données fournissent déjà une partie de ces informations. Toutefois, les métadonnées actuellement fournies sont plus des descriptions techniques des jeux de données que des informations compréhensibles destinées aux usagers (Timpf *et al.*, 1996; Harvey, 1998). Les métadonnées sont donc dans la pratique très peu utilisées, laissant les utilisateurs experts comme non-experts en géomatique dans un état d'ignorance concernant la qualité des données géospatiales qu'ils manipulent.

De plus, les métadonnées sont très rarement assez détaillées (Hunter, 2001; Gan et Shi, 2002), étant la plupart du temps une description au niveau du jeu de données. La qualité décrite dans les métadonnées est alors une agrégation de qualités hétérogènes des objets composant le jeu de données (ex : « la précision spatiale du jeu de données varie entre 10 m et 1 km »). Pour être utiles, celles-ci devraient décrire les données à un niveau de détail plus fin, comme au niveau de l'instance d'objet ou même de la valeur d'un attribut, ce qui fournirait un grand volume

d'information aux utilisateurs. Toutefois, les humains ne résolvent pas les situations complexes avec un grand volume d'information, mais en sélectionnant les informations pertinentes (Klein, 1999). Dans ce sens, Fisher (2001) mentionne que « le défi dans un monde riche en information n'est pas seulement de rendre l'information disponible aux personnes en toute place et sous toutes les formes, mais de réduire la surcharge d'information en rendant l'information pertinente pour la tâche effectuée et en fonction du bagage de connaissances présumé des utilisateurs » (traduction libre).

Ainsi, afin de fournir uniquement des informations pertinentes aux usagers au niveau de détail dont ils ont besoin, il existe un besoin pour différentes vues agrégées de ces informations de qualité, contextuelles aux profils des utilisateurs et à la tâche qu'ils effectuent.

Ce contexte n'est pas exclusif aux données géospatiales. Par exemple, les gestionnaires d'entreprises ont également besoin d'une vue agrégée des informations décrivant leur compagnie. Les gestionnaires de grandes chaînes de magasins ne sont en général pas intéressés par des listes de toutes les ventes faites dans leurs magasins, mais par des agrégations des ventes par type de produit, intervalle de temps, région, etc. Dans ces domaines, les décideurs utilisent maintenant des outils du domaine du *Business Intelligence* les aidant dans leurs tâches. Parmi ces outils, les tableaux de bord exécutifs fournissent des informations agrégées, nommées indicateurs, sur différents aspects des organisations.

L'objectif de cet article est de présenter une approche basée sur des indicateurs permettant de communiquer l'information relative à la qualité des données géospatiales aux utilisateurs. Étant donné que les problèmes de qualité deviennent rapidement complexes (ex. diversité des paramètres décrivant la qualité, granularité des informations sur la qualité, hétérogénéité spatiale), les utilisateurs ciblés dans cet article sont des experts ayant une bonne connaissance de la géomatique. En effet, dans le contexte actuel, même les experts en géomatique ont beaucoup de difficulté à se prononcer sur la qualité des données pour une application précise dans un secteur précis. Les données utilisées résultent souvent de l'intégration de différentes sources pouvant avoir été collectées suivant différentes normes, à différentes époques avec des technologies diverses. Ainsi, les données manipulées dans les SIG sont souvent très hétérogènes et l'adéquation de ces données à une application spécifique demeure complexe. Bien que l'application présentée vise des usagers experts en géomatique, l'approche globale est en partie applicable à des utilisateurs non-experts. Cette approche fournit des informations pertinentes aux utilisateurs relativement à la qualité des données qu'ils manipulent afin de réduire les risques de mauvaises utilisations de ces données. L'utilisation de logiciels SIG dans des processus de prise de décision sera abordée dans la première section. De manière plus spécifique, nous présenterons les SIG comme un processus de communication entre des producteurs et des utilisateurs de données. Nous présenterons l'incertitude reliée aux processus de prise de décision et mentionnerons les informations relatives à la qualité des données actuellement communiquées aux

utilisateurs et leurs limites pour le support à la prise de décision. La deuxième section présente les concepts et caractéristiques des tableaux de bord de gestion et des indicateurs. Les caractéristiques des tableaux de bord et indicateurs dans le contexte géospatial seront présentées. Finalement, la troisième section présente un aperçu d'un prototype intégré dans une interface cartographique, permettant de gérer et communiquer ces indicateurs.

2. SIG et prise de décision

2.1. *Les SIG – un processus de communication*

Shannon (1948) définit la communication comme « reproduire en un point exactement ou approximativement un message sélectionné en un autre point » (traduction libre). Basé sur les adaptations de la théorie de la communication de Shannon pour le domaine de la communication de masse (ex : journalisme) et pour les sciences cognitives (ex : perception, interprétation de signaux), Bédard (1986) identifie les SIG (en tant que système organisationnel) comme étant un processus de communication complexe entre des producteurs et des utilisateurs de données géospatiales. Afin de prendre une décision, les personnes perçoivent des signaux du monde réel, les interprètent, et procèdent à une abstraction afin de générer un modèle cognitif servant à cette prise de décision. Les signaux perçus peuvent provenir soit d'une observation directe de la réalité, soit d'une autre personne (ou machine) mandatée pour communiquer une information. Dans le cas des utilisateurs de logiciel SIG, les signaux perçus proviennent presque toujours d'un observateur autre que l'utilisateur, créant ainsi un processus de communication entre l'observateur de la réalité (ex : géomètre, forestier, géologue) et l'utilisateur du logiciel SIG. De nos jours, il est même de plus en plus fréquent pour un utilisateur de logiciel SIG d'utiliser des données multisources.

Une caractéristique importante des processus de communication est le besoin de connaissances communes (identifié en anglais par le concept de *commonness*) entre producteurs de signaux et récepteurs (pouvant être des individus ou des machines) (Shannon, 1948; Bédard, 1986; Martinet et Marti, 2001). L'ensemble des connaissances d'un agent est identifié comme étant son cadre de référence. Plus les *connaissances communes* sont importantes entre le producteur et l'utilisateur d'une information, plus les risques de distorsion du message seront faibles. En pratique, cette communication est toujours imparfaite à cause des différences entre sources et cibles. Afin de faciliter la communication entre les agents, Martinet et Marti préconisent l'utilisation d'un langage le plus proche possible de la cible. Les SIG communiquent donc toujours les informations avec un certain biais, mais l'emploi d'un langage graphique proche des connaissances des utilisateurs des données peut limiter ce biais.

2.2. *Prise de décision et incertitude*

Les utilisateurs de SIG manipulent les données géospatiales afin d'obtenir des informations pouvant être utilisées dans un processus de prise de décision plus large (ex : prendre le chemin le plus court pour se rendre quelque part, trouver la parcelle cadastrale idéale pour construire un bâtiment). Mintzberg (1979) définit la décision comme « le signal d'une intention explicite d'agir » (traduction libre). La décision ne se limite pas à l'action. Fernandez (2000) identifie quatre étapes dans un processus de prise de décision, soit (1) la formalisation du désir, lorsque l'agent prend conscience de la situation, (2) l'instruction, lorsque l'agent collecte les informations, analyse des situations précédentes et des solutions potentielles, (3) le choix, lorsque l'agent identifie l'action à effectuer et évalue ses limites et enfin (4) l'action. Il formule également plusieurs conditions pour prendre une bonne décision. Les décisions sont prises (1) afin d'atteindre un objectif, (2) selon la situation perçue, (3) selon l'expérience et le référentiel de valeurs du décideur, (4) selon ses motivations, (5) en fonction de la mesure des risques et (6) selon les moyens conférés et disponibles. Basé sur des observations pratiques de différents types de décideurs, Klein (1999) affirme que l'intuition et les simulations mentales sont centrales dans la prise de décision, basées respectivement sur l'expérience et l'imagination. Il explique que « l'intuition dépend de l'usage de l'expérience pour reconnaître des patrons clés indiquant la dynamique de la situation » (Klein, 1999; p. 31 – traduction libre). Le modèle RPD de Klein (*Recognition-Primed Decision*) offre un cadre théorique pour les processus de prise de décision. Ce modèle souligne l'importance des indices pertinents qui aident les décideurs à reconnaître une situation, évitant une surcharge possible d'information.

Fernandez (2000) différencie la décision du calcul. D'un côté, le calcul permet le choix d'une bonne solution rationnelle et est automatisable. De l'autre, la décision est basée sur des informations incertaines, imprécises et insuffisantes, mettant en jeu le contexte, les acteurs et la situation. Cette incertitude peut apparaître à différents niveaux. Le concept d'incertitude, ainsi que d'autres termes liés au domaine de la qualité, est présenté, entre autres, plus précisément dans un article de Fisher (1999).

Quand une personne fait face à de l'incertitude lors d'un processus de prise de décision et est consciente du type d'incertitude et de son importance, il peut choisir entre (1) ne rien faire, (2) essayer de réduire cette incertitude ou (3) prendre la décision et accepter les conséquences possibles, « absorbant » ainsi cette incertitude (Bédard, 1986). Epstein *et al.* (1998) suggèrent de réduire l'incertitude en (1) obtenant plus d'information et/ou (2) améliorant la qualité de l'information disponible. L'incertitude résiduelle alors absorbée est le niveau de risque relié à l'utilisation de cette information (Bédard, 1986; Epstein *et al.*, 1998). Le niveau de risque acceptable dépend du décideur, de l'application ou du contexte institutionnel.

Les décisions sont donc toujours basées sur des informations incertaines et incomplètes. Les décideurs ont alors le choix entre prendre la décision en acceptant l'incertitude résiduelle ou collecter de nouvelles informations pour diminuer cette

incertitude. Ceux-ci utilisent des indices (ou indicateurs) afin de caractériser une situation, diminuer l'incertitude et donc orienter leur décision.

2.3. *Communication de l'information sur la qualité des données géospatiales*

Les producteurs de données fournissent de plus en plus souvent des métadonnées documentant différents aspects des jeux de données, afin de renseigner les utilisateurs sur les caractéristiques des données qu'ils utilisent. Selon les principales normes en géomatique, les métadonnées devraient fournir de l'information relative à la qualité des données géospatiales, telle que la précision spatiale, la complétude (omission, commission) ou la consistance logique de la base de données (Guptill et Morrison, 1995; FGDC, 2000; ISO-TC/211, 2003).

Toutefois, l'expérience montre que ces métadonnées sont complexes à comprendre et à utiliser pour des utilisateurs non-experts mais aussi par les experts en données géospatiales, restant de ce fait la plupart du temps inutilisées (Timpf *et al*., 1996; Frank, 1998). Ce sont plus des descriptions techniques dont le contenu découle des procédures de production des jeux de données, que des informations compréhensibles et pertinentes pouvant être utilisées par des utilisateurs de données pour supporter leur processus de prise de décision (Frank, 1998).

De plus, les métadonnées fournissent la plupart du temps une description des données au niveau du jeu de données. Comme la qualité peut être très hétérogène dans l'espace et dans le temps, des métadonnées à un niveau de détail plus fin, tel que l'occurrence d'objet ou l'attribut, seraient souvent nécessaires (Hunter, 2001; Gan et Shi, 2002).

Les métadonnées ne sont donc pas un moyen efficace de communiquer les informations relatives à la qualité des données aux utilisateurs de données. Différentes approches peuvent alors être explorées pour aborder ce problème.

Basé sur le paradigme de communication des SIG développé par Bédard, la figure 1 illustre de façon théorique quelques solutions possibles, en décrivant les relations entre les connaissances des producteurs et utilisateurs de données ainsi que la position des métadonnées dans ce cadre de référence. Les cercles représentent les cadres de référence des producteurs et des utilisateurs de données et leurs intersections correspondent aux connaissances communes.

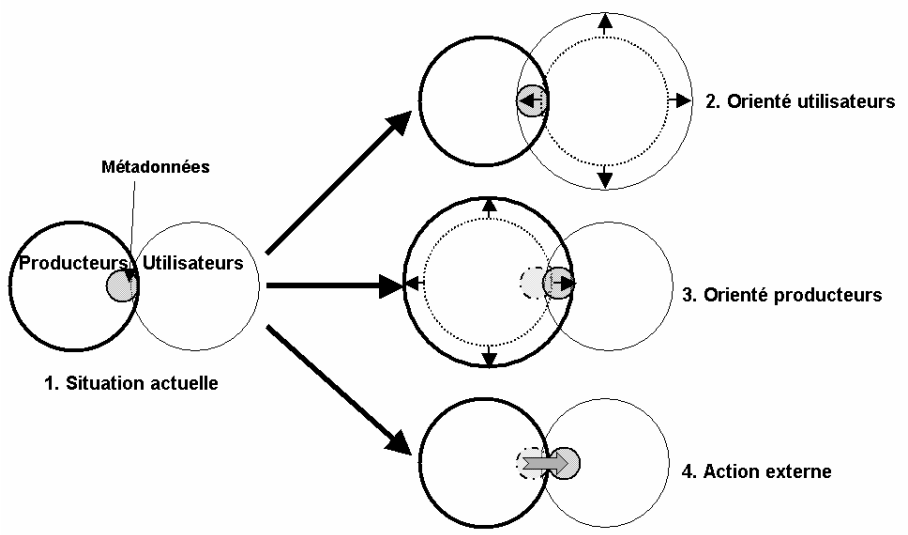

Figure 1. *Les métadonnées dans le processus de communication utilisateurs-producteurs.*

1. Situation actuelle : Actuellement, les producteurs de données et les utilisateurs ont des connaissances communes faibles concernant les données géospatiales. Les métadonnées utilisent un vocabulaire technique, et sont donc situées dans le cadre de référence des producteurs de données. Les utilisateurs n'ayant généralement pas de connaissance dans le domaine de la géomatique (projections, échelles, techniques d'acquisition et de traitement des données, etc.), ils ont une compréhension très réduite des métadonnées.

2. Approche orientée utilisateurs : Une solution possible serait d'accroître le cadre de référence des utilisateurs, en leur « enseignant » des concepts et du vocabulaire relié aux données géospatiales. Ceci augmenterait la connaissance des utilisateurs, accroissant leur expertise et donc les connaissances communes entre producteurs et utilisateurs. Toutefois, cela signifie faire des experts d'utilisateurs non-experts.

3. Approche orientée producteurs : L'inverse serait d'accroître le cadre de référence des producteurs – *e.g.* essayer de vulgariser le vocabulaire technique des métadonnées– accroissant ainsi les connaissances communes. Certaines normes

recommandent une telle approche en encourageant l'utilisation de vocabulaire non technique et les descriptions en texte libre afin de rendre plus accessible la compréhension des métadonnées. Toutefois, cela représente un travail additionnel de la part des producteurs de données et demeure souvent insatisfaisant en regard de la compréhension des utilisateurs.

Ce problème est semblable à des problèmes de communication classiques entre deux individus ne parlant pas la langue de l'autre (ex *:* un chinois et un français). Les cas 2 et 3 signifient qu'une des deux personnes apprend le langage de l'autre. Toutefois, comme cela requière des ressources considérables (temps et souvent argent), il est fréquent de faire appel à une troisième personne connaissant les deux langages pour traduire le message.

4. Action externe : Ce cas, en linguistique, correspond au rôle du traducteur qui est chargé de faire le lien entre les deux agents. Le message produit par l'un des agents (*i.e.* producteur de données) dans un langage (*i.e.* métadonnées techniques) est traduit dans le langage de l'autre agent (*i.e.* utilisateur des données). Comme tout processus de communication, la distorsion du message est minimisée mais souvent inévitable. Ce dernier cas est l'approche examinée dans cet article, traduisant des métadonnées stockées en général dans des fichiers texte externes aux données, en des indicateurs plus facilement interprétables intégrés dans l'interface du SIG. L'approche présentée va également plus loin en effectuant cette traduction de manière contextuelle, en offrant aux usagers un accès hiérarchique aux indicateurs et en l'avertissant potentiellement de risques de mauvaises utilisations.

3. Tableaux de bord et indicateurs pour supporter la prise de décision

3.1. *Tableaux de bord*

L'analogie avec les tableaux de bord automobiles peut illustrer la façon de représenter une réalité complexe en utilisant un modèle simplifié. Le tableau de bord d'une voiture permet au conducteur d'obtenir en temps réel des informations limitées, mais souvent suffisantes, concernant le système plus complexe qu'est son véhicule. Même si le tableau de bord donne une vision incomplète et souvent imprécise de la réalité, cela permet au conducteur de prendre des décisions rapides telles que diminuer sa vitesse, ajouter de l'huile ou s'arrêter prendre de l'essence.

Un tableau de bord de gestion est défini par Voyer comme « une façon de sélectionner, d'agencer et de représenter des indicateurs essentiels et pertinents, de façon sommaire et ciblée […] fournissant à la fois une vision globale et la possibilité de forer dans les niveaux de détail. » (Voyer, 2000; p.39). Dans le domaine des systèmes de support à la prise de décision, les tableaux de bord de décision (*executive dashboard*) sont aussi nommés *scorecard, balanced scorecard, scoreboard, steering panel* ou *control panel*. Les tableaux de bord se concentrent surtout sur la qualité de l'information et non sur sa quantité. Ils représentent les

indicateurs de façon compréhensible, suggestive et attractive afin de faciliter leur visualisation. Ils présentent un aperçu représentatif de la situation, permettant ensuite d'accéder aux données plus détaillées au besoin. Le tableau de bord doit être contextuel, le décideur pouvant sélectionner ses propres indicateurs, avec la représentation qu'il préfère, afin de produire son tableau de bord personnel.

De nombreuses organisations utilisent des tableaux de bord. Par exemple, le gouvernement canadien encourage l'utilisation d'indicateurs dans son administration. Les grandes compagnies comme les banques, compagnies d'assurances utilisent des indicateurs. Les grands organismes internationaux (banque mondiale, Nations Unies, agences américaine et canadienne de développement international, etc.) utilisent également des indicateurs sociaux, économiques, géopolitiques ou environnementaux.

Les tableaux de bord permettent la visualisation d'un ensemble d'indicateurs. En effet, l'utilisation d'un seul indicateur serait trop dangereuse (Kaplan et Norton, 1992). Prenez par exemple un pilote d'avion qui a besoin d'information sur de nombreuses variables telles que l'essence, l'altitude, la vitesse de l'air, la position, la destination, etc. Ces informations ne peuvent pas être fournies par un seul indicateur. Le nombre d'indicateur doit cependant être limité afin d'éviter une surcharge d'information. D'après Miller (1956), l'être humain peut percevoir 7±2 éléments en même temps. Ce « nombre magique » est maintenant largement utilisé pour la communication d'informations et peut donc être utilisé dans la conception de tableaux de bord en géomatique.

3.2. *Indicateurs*

Le *Jackson Community Council* (Plan Canada, 1999) définit un indicateur comme « une manière de voir un portrait général en regardant un petit morceau de celui-ci » (traduction libre). Fernandez (2000) le définit comme « une information ou un regroupement d'informations contribuant à l'appréciation générale d'une situation par le décideur » (p.232). Klein (1999) identifie les indicateurs comme étant des indices situés au centre des processus de prise de décision en supportant les intuitions des décideurs. L'objectif d'un indicateur est de mesurer une situation et d'initier une réaction du décideur, la réaction pouvant être de ne rien faire.

Le système doit fournir un ensemble d'indicateurs que les usagers peuvent adapter à leur contexte si besoin ou des indicateurs pouvant être partagés à l'intérieur d'une même communauté d'utilisateurs. Il devrait aussi permettre aux décideurs de créer leurs propres indicateurs et règles pour les calculer.

La valeur d'un indicateur peut être basée sur une donnée unique ou résulter d'un calcul impliquant plusieurs données. Ces données doivent être techniquement accessibles. Elles peuvent être déjà disponibles dans une base de données accessible ou provenir d'autres sources, tels que des opinions d'experts ou de collègues.

Comme les données sont valides pour une certaine durée dans le temps (*life time*), leur actualité doit être prise en considération.

Les caractéristiques des indicateurs peuvent être décrites sur une feuille d'indicateur que les utilisateurs peuvent consulter et modifier si nécessaire. Cette feuille peut fournir par exemple de l'information sur la définition de l'indicateur, ses représentations possibles, sur les considérations reliées à son utilisation et interprétation, ses mécanismes de validation, etc. Les indicateurs peuvent représenter différents types d'information, tant quantitatifs que qualitatifs. Il est préférable de fournir des indicateurs « flous », *i.e* de précision limitée (Fernandez, 2000), tel qu'un intervalle de valeurs ou une échelle qualitative, car des valeurs trop précises encourageraient l'utilisateur à se concentrer sur la valeur et pas sa signification dans une perspective globale.

Diverses représentations peuvent être utilisées pour visualiser la valeur d'un indicateur, telles que des nombres, symboles, icônes, pictogrammes, tables, graphiques, textes, sons, images, etc. Il est également possible d'utiliser des fenêtres *pop-up*, alarmes visuelles ou sonores, etc., qui sont souvent des façons efficaces de capter l'attention des utilisateurs afin qu'ils se concentrent sur l'essentiel.

4. Tableaux de bord et indicateurs pour la prise de décision géospatiale

4.1. Tableaux de bord et système MUM

Les tableaux de bord de gestion font partie du domaine du support à la prise de décision et de manière plus spécifique du *Business Intelligence*. Certains travaux ont été faits pour adapter des outils du *Business Intelligence* dans le domaine de la géomatique, tels que pour le *Data Mining* Spatial, le SOLAP (*Spatial On-Line Analytical Processing*) et les entrepôts de données géospatiales (Miller et Han, 2001; Rivest *et al.*, 2001). Plusieurs logiciels développés pour le domaine du *Business Intelligence* visent à créer et maintenir des tableaux de bord de gestion. Ils sont par exemple Esperant et Media de Speedware, Metrics Manager de Cognos, EIS de SAS, Oracle Balanced Scorecard de Oracle, Hyperion Performance Scorecard de Hyperion, Crystal Application de Crystal Decisions. Comme les tableaux de bord fournissent habituellement des informations à différents niveaux de détail, la plupart des systèmes reposent sur des bases de données multidimensionnelles. Une telle structure conçue pour la gestion des informations relatives à la qualité des données géospatiales est décrite dans Devillers *et al.* (2004) et est utilisée pour la conception du prototype de tableau de bord géospatial. Cette structure permet de gérer les informations de qualité à différents niveaux de détails.

Le tableau de bord, associé à un SIG devrait être capable de :

- *Représenter l'information de qualité sous la forme d'indicateurs* : les indicateurs fournissent des informations brutes ou agrégées sur la qualité des données géospatiales. Les indicateurs doivent être présentés sur un tableau de bord faisant partie intégrante de l'interface du SIG et peuvent être rendus visibles ou non selon le désir des utilisateurs;

- *Fournir des indicateurs en temps réel :* étant donné que les utilisateurs peuvent vouloir ajouter ou retirer des données dans leurs SIG, modifier leurs profils personnels (ex : tolérance face au risque), etc., les valeurs des indicateurs doivent être recalculées à chaque modification du contexte de l'utilisateur.

- *Fournir des indicateurs en fonction de l'étendue spatiale visualisée* : la qualité peut être très hétérogène dans l'espace et dans le temps. Par exemple, un secteur d'une carte a pu être mis à jour récemment avec une grande précision et exactitude tandis qu'un autre secteur de la même carte présente des données anciennes et imprécises. Les valeurs des indicateurs doivent donc être calculées à partir des qualités des objets situés dans la zone visualisée par l'utilisateur, et non pas uniquement représenter la qualité moyenne de l'ensemble des données du jeu de données. L'utilisateur doit également pouvoir obtenir la qualité moyenne d'une zone qu'il définit de façon ad hoc, soit en lui permettant de tracer lui-même cette zone (ex : création d'un polygone), soit en lui offrant une liste de zones prédéfinies (ex : villes, quartiers). Cela implique une mise à jour des indicateurs lorsque l'utilisateur navigue dans sa vue (ex : *Zoom in, Zoom out*, *Pan*);

- *Permettre aux utilisateurs de sélectionner les indicateurs pertinents dans leur contexte ou définir leurs propres indicateurs* : différents utilisateurs ont des profils, objectifs et intérêts différents. Un ensemble d'indicateurs prédéfinis doit être mis à la disposition des utilisateurs. Toutefois, les utilisateurs doivent être capables de voir comment sont calculés ces indicateurs, de modifier ces procédures et si possible de permettre la création de nouveaux indicateurs;

- *Permettre aux utilisateurs de visualiser les indicateurs à différents niveaux de détail* : les indicateurs doivent être organisés de manière hiérarchique (indicateurs et sous-indicateurs) afin d'éviter aux utilisateurs une surcharge d'information. On conserve ainsi un nombre d'indicateurs conforme à la loi de Miller (7±2 indicateurs) tout en permettant aux utilisateurs d'approfondir l'exploration des informations de qualité de manière intuitive;

- *Permettre aux utilisateurs de mettre des poids sur les différents indicateurs, en fonction de leur importance dans le contexte d'utilisation des données* : certains indicateurs peuvent avoir plus d'importance que d'autres. Par exemple, la complétude des données peut être beaucoup plus importante que la précision temporelle pour certaines applications. Ces poids entrent en jeu lors de l'agrégation des sous-indicateurs en indicateurs de plus hauts niveaux ;

- *Permettre la définition et la gestion des profils des utilisateurs (niveau de risque acceptable, etc.)* : différents utilisateurs peuvent vouloir différentes façons

d'agréger des indicateurs. Par exemple, certaines personnes peuvent avoir une plus grande tolérance face aux risques que d'autres dans leurs décisions, dépendant par exemple de leur contexte organisationnel (ex : une personne utilisant un SIG pour planifier une sortie loisir en famille pourra accepter plus de risques qu'un gestionnaire utilisant un SIG pour gérer des épidémies dans un organisme de santé environnementale);

- *Offrir différentes représentations des indicateurs que les utilisateurs peuvent sélectionner* : certains utilisateurs peuvent préférer certaines représentations pour les indicateurs (ex : feux de circulation, histogrammes, compteur de vitesse). Les utilisateurs doivent pouvoir choisir le mode de représentation qu'ils préfèrent parmi un choix de représentations dépendant du type d'indicateur et des valeurs qu'il communique (ex : quantitatif, qualitatif);

- *Offrir un mode de visualisation cartographique des indicateurs de qualité* : en plus d'une représentation des indicateurs dans un tableau de bord, la valeur des indicateurs doit pouvoir être représentée sur la carte. Par exemple, un indicateur ayant une représentation du type feu de circulation (vert/jaune/rouge) pourra avoir une valeur jaune représentant l'ensemble des données visualisées dans l'interface. L'utilisateur pourra passer en mode de représentation cartographique de la qualité et ainsi avoir une meilleure idée des qualités individuelles des objets (ex : précision spatiale), chaque objet affiché étant représenté en vert, jaune ou rouge, dépendamment de sa qualité. Cette représentation permet entre autres d'identifier rapidement l'hétérogénéité spatiale de la qualité.

- *Activer des alarmes automatiquement lorsque certaines conditions sont atteintes* : des signaux sonores ou visuels peuvent être émis pour capter l'attention des utilisateurs à certains moments critiques, comme lorsqu'un indicateur dépasse la tolérance définie par l'utilisateur.

La figure 2 présente un schéma général de la création des indicateurs allant être affichés dans le tableau de bord du système MUM. (1) Une interface permet de collecter les informations caractérisant l'utilisateur (contexte, style de gestion, etc.) et conserve ces informations. (2) Une base de données d'indicateurs prédéfinis permet à l'utilisateur de sélectionner et éventuellement modifier des indicateurs existants. L'utilisateur peut aussi définir et stocker de nouveaux indicateurs. Cette étape permet aux utilisateurs de personnaliser leurs indicateurs et leur tableau de bord. (3) Les métadonnées et autres informations pertinentes décrivant les jeux de données sont intégrées et structurées dans une même base de données à différents niveaux de détails. Ce processus d'intégration doit idéalement être automatique ou semi-automatique afin d'assurer une certaine flexibilité au système. (4) Les indicateurs ayant été sélectionnés, leurs valeurs sont calculées en utilisant la règle d'agrégation définie, celle-ci dépendant de l'indicateur, de l'information disponible pour le calculer et du profil de l'usager. (5) Les indicateurs sélectionnés sont alors affichés dans l'interface du SIG selon le mode de représentation choisi par l'usager afin d'informer l'utilisateur de la qualité des données qu'il utilise. Ces indicateurs

sont par la suite mis à jour dès que des changements ont lieu (ex : changement au profil de l'utilisateur, navigation dans l'interface cartographique, navigation à l'aide de fonctions OLAP).

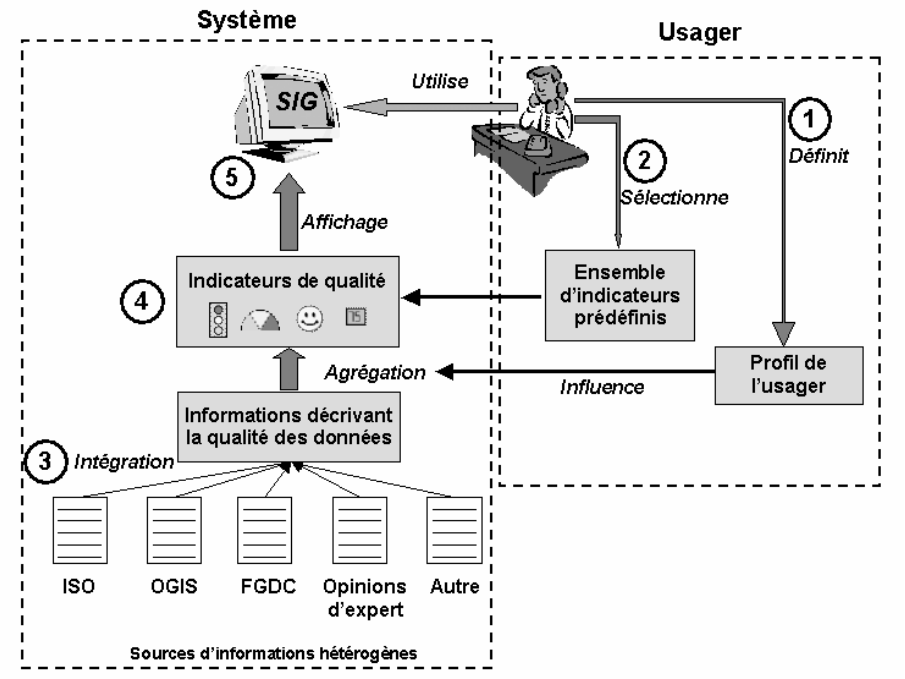

Figure 2. *Fonctionnement simplifié du système MUM.*

4.2. *Indicateurs de qualité des données géospatiales*

Les mauvaises utilisations de données géospatiales peuvent apparaître à différents niveaux lors de l'utilisation d'un SIG. Par exemple, un utilisateur peut faire une mauvaise utilisation d'une fonction d'un SIG (ex : interpoler des données nominales de points ou afficher au 1 :10 000 une carte créée à l'échelle 1 :1 000 000). Un utilisateur peut également obtenir un mauvais résultat de fonctions dans un SIG lorsque les données contiennent des erreurs (ex : mesures de distances très précises basées sur des données largement inexactes ou encore

calculer un nombre d'objets alors que la complétude du jeu de données est médiocre). Il est également nécessaire de prendre en considération que beaucoup d'utilisateurs font appel aux SIG pour visualiser les données, sans forcément utiliser de fonctions d'analyse. Donc, les outils visant à réduire les risques de mauvaises utilisations de données géospatiales doivent se concentrer à la fois sur les erreurs issues de la manipulation des opérateurs d'un SIG (ex : opérateurs topologiques et métriques) et sur la mauvaise interprétation de données affichées par le SIG. Nous pouvons ainsi identifier deux types d'avertissements pouvant réduire les risques de mauvaises utilisations :

– Avertissements de manipulation :

- Messages d'opérations illogiques (Hunter et Reinke, 2000) : des avertissements sonores ou visuels peuvent être communiqués aux utilisateurs lorsqu'une manipulation pouvant engendrer un risque est effectuée sur des données dans le SIG (ex : requêtes, zoom, mise à jour, etc.) (cf. figure 3). Hunter et Reinke donnent plusieurs exemples d'opérations illogiques pouvant être traduites en algorithmes tels que :

```
IF command_name = 'calculate_map_distance'
    AND map_units = null
    OR distance_units = null
    OR projection_type = null
THEN generate map_distance_warning
```

De telles règles pourraient limiter les risques les plus courants de manipulation en émettant des avertissements ou en désactivant certaines fonctions du SIG pouvant induire un risque. Une connaissance des données, provenant par exemple des métadonnées (ex : exactitude des données) ou directement de la structure des données (ex : précision numérique des données), est nécessaire. Les règles doivent être définies par des experts puis stockées dans une base de règles pouvant être interrogée par le système lors de chaque opération effectuée dans le SIG.

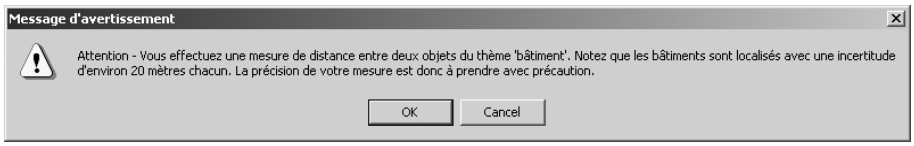

Figure 3. *Exemple de message d'opération illogique*

– Avertissements d'état :

- Indicateurs de statut : un indicateur de statut fournit une information « brute » décrivant une ou plusieurs caractéristiques internes des données. Par exemple, si la précision spatiale est de 13 mètres, la valeur de l'indicateur « précision spatiale » pourrait être par exemple « 13 mètres » ou un intervalle tel que « 10-20 mètres »;

- Indicateurs de risque : les indicateurs de risque fournissent une information « relative », provenant de la comparaison entre des caractéristiques des données (données brutes) et les besoins des utilisateurs, la divergence étant exprimée en terme de niveau de risque. Ceux-ci représentent des informations sur la qualité des données, la qualité étant définie comme l'adéquation à l'usage (*fitness for use*). Par exemple, si la précision spatiale des données est de « 13 mètres » et l'usager désire des données avec une précision de « 1 mètre », la valeur de l'indicateur précision spatiale pourrait être affichée avec une lumière rouge, avertissant l'utilisateur de l'importance de la différence. L'utilisateur aurait alors une idée du risque relié à l'utilisation des données pour ce critère de qualité. Le calcul de ce type d'indicateur implique une qualification de données quantitatives (passer de « 13 mètres » à une lumière rouge dans cet exemple). Cette qualification de l'information est complexe et peut être faite de différentes façons. Différents modes de représentation peuvent être utilisés pour les indicateurs de risque, tels que des feux de circulation, des compteurs de vitesse, des *smiley*, etc. Ces indicateurs généralement binaires ou trinaires permettent de représenter un indicateur passant un message du type *go/no go* ou mauvais/moyen/bon.

Les utilisateurs doivent avoir accès à des descriptions des indicateurs proposés. Un exemple de feuille descriptive d'indicateurs est présenté sur la figure 4. Cette feuille permet la description de différents aspects de l'indicateur tels que :

- Définition/signification de l'indicateur;
- Méthode utilisée pour calculer la valeur de l'indicateur;
- Mode de représentation (ex : valeur simple, feux de circulation, *smiley*);
- Importance de l'indicateur pour l'utilisateur;

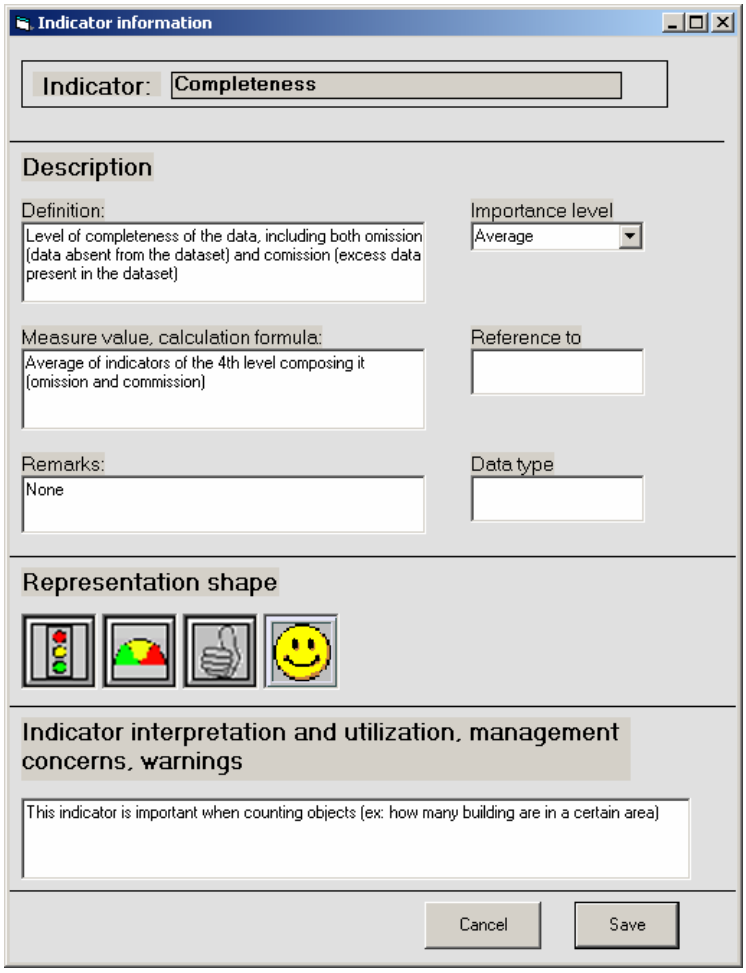

Figure 4. *Exemple de fiche descriptive d'un indicateur de qualité.*

4.3. *Prototype du système MUM*

Un prototype du système MUM (Manuel à l'Usager Multidimensionnel) a été développé afin de tester l'approche de communication de la qualité sous la forme d'indicateurs. Le prototype a été programmé en orienté-objet, utilisant des objets de différentes applications, et se base principalement sur trois technologies : SQL

Server, GeoMedia et Proclarity. Une base de données multidimensionnelle gérant les informations de qualité a été implantée avec le serveur OLAP SQL Server/ Analysis Services de Microsoft. Le modèle de données utilisé est décrit dans Devillers *et al.* (2004). Les fonctionnalités cartographiques du prototype (*zoom in*, *out*, cartes thématiques, etc.) ont été développées avec des objets du logiciel GeoMedia Professional 5 d'Intergraph. Les fonctionnalités OLAP, permettant à l'utilisateur de naviguer dans une base de données multidimensionnelle, ont utilisé des objets du logiciel OLAP-client Proclarity 5. Les données utilisées dans le prototype sont un extrait de la Base Nationale de Données Topographiques du Canada (BNDT) pour le secteur de la ville de Sherbrooke (Québec, Canada). Ces données incluent les routes, bâtiments principaux, rivières, etc. pour des zones de qualité variable.

Pour tester le prototype, les indicateurs proposés par le système sont principalement basés sur la norme internationale ISO 19113 (Principes de qualité) et 19115 (Métadonnées). Les indicateurs de qualité sont gérés de façon hiérarchique selon une dimension dans la base de données multidimensionnelle (cf. Devillers, 2004). Les indicateurs détaillés sont basés sur une ou plusieurs métadonnées et ceux de plus hauts niveaux sont des agrégations des indicateurs les composant. Pour le prototype, seuls des indicateurs de risque ont été créés, la qualité étant communiquée sous une forme qualitative utilisant différentes représentations telles que des feux de circulation (vert/orange/rouge), *smiley*, etc.

Le prototype permet différentes fonctionnalités telles que :

- La sélection par l'usager d'indicateurs prédéfinis, stockés hiérarchiquement dans une base de données MS-Access. Les indicateurs sélectionnés sont alors affichés dans le tableau de bord;
- La définition d'un profil minimal de l'utilisateur incluant entre autres sa tolérance face au risque et les indicateurs qu'il a sélectionné;
- La visualisation de fiches descriptives pour chacun des indicateurs présentant leur définition, type de représentation, mode de calcul, etc. (cf. figure 4);
- La visualisation des indicateurs dans un tableau de bord pouvant inclure jusqu'à 9 indicateurs plus un indicateur global (cf. figure 5). Ces indicateurs ont été sélectionnés par l'utilisateur parmi une liste hiérarchique d'indicateurs prédéfinis. L'indicateur global représente une agrégation des valeurs des indicateurs sélectionnés, la méthode d'agrégation (ex : maximum, moyenne) dépendant du profil défini par l'utilisateur. L'indicateur global présente une vue générale de la concordance entre la qualité interne des données et les besoins exprimés par les utilisateurs. Utilisant une symbologie de type feu de circulation, une lumière verte signifie qu'il peut manipuler les données sans risque apparent. Des lumières oranges ou rouges l'encouragent à explorer les indicateurs le composant;

- La visualisation cartographique des indicateurs, les valeurs de qualité étant associées à chaque entité géométrique (cf. figure 5). L'utilisateur doit identifier l'indicateur de qualité qu'il désire représenter et chaque objet de la carte prend alors la valeur de qualité lui étant associée (carte thématique de la qualité utilisant les couleurs vert/orange/rouge);

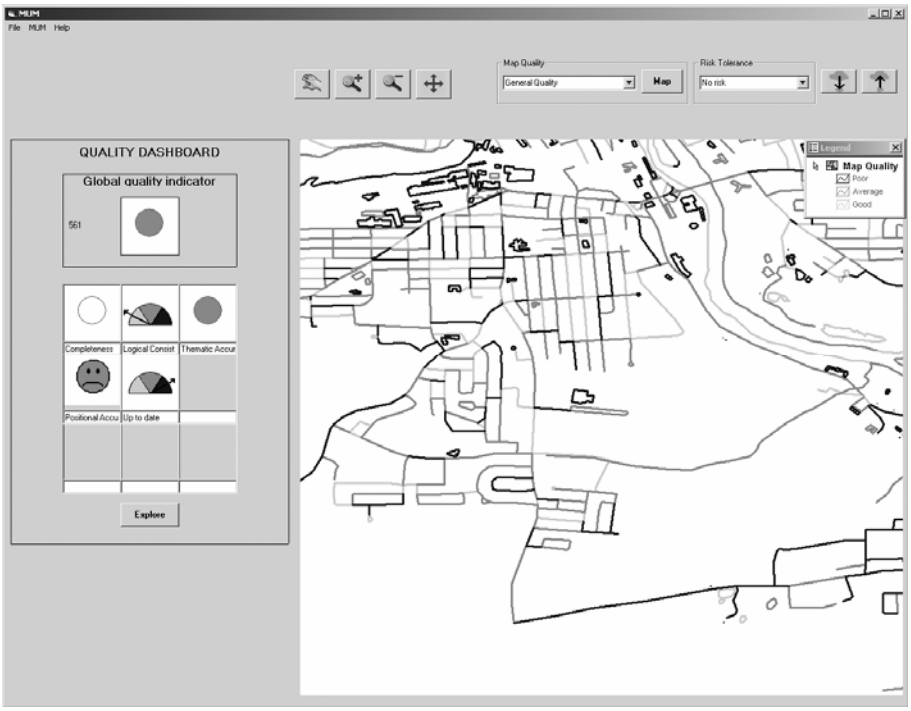

Figure 5. *Interface cartographique du MUM avec tableau de bord et indicateurs (gauche) et représentation cartographique de la qualité (droite). La symbologie vert/jaune/rouge est représentée ici par des niveaux de gris (de gris clair à foncé respectivement).*

- La possibilité pour l'utilisateur d'utiliser des fonctions de type OLAP telles que Drill-Down et Roll-Up afin de naviguer dans les données multidimensionnelles à différents niveaux de détail (ex : visualiser la qualité globale du jeu de données, puis la qualité des routes uniquement, et enfin la

qualité d'une seule route). Ces outils permettent également de visualiser les indicateurs de qualité à différents niveaux de détail à l'intérieur de la hiérarchie d'indicateurs (cf. figure 6). La figure 6 présente un indicateur et les sous-indicateurs le composant. L'utilisateur peut utiliser les opérateurs de forage OLAP afin de visualiser un niveau plus détaillé ou plus général de la hiérarchie;

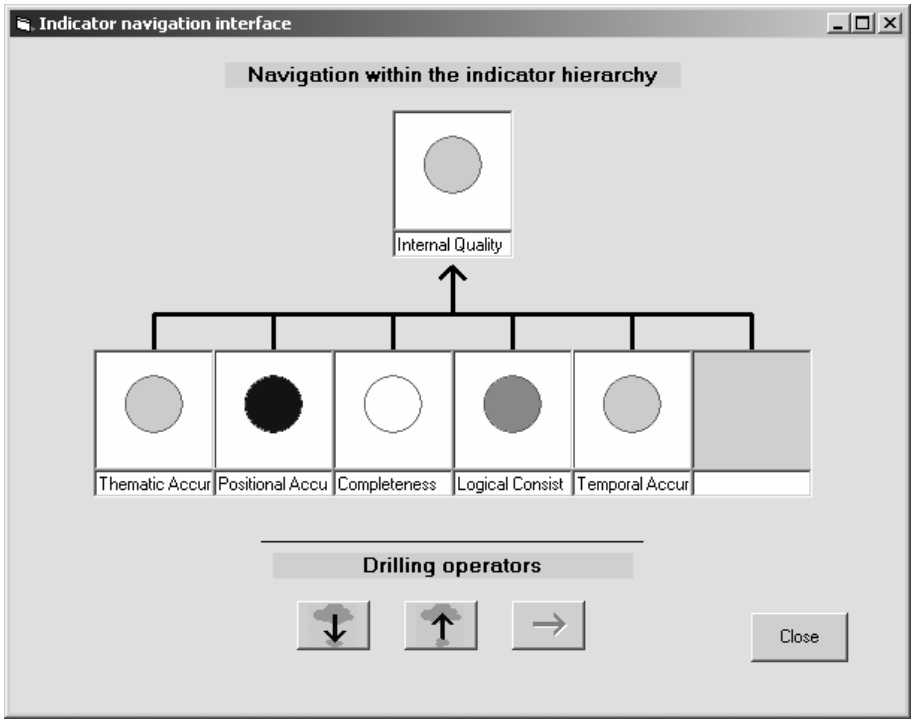

Figure 6. *Outil permettant la navigation dans la hiérarchie d'indicateurs de qualité.*

Le tableau de bord permet aux utilisateurs de naviguer dans la hiérarchie des indicateurs de qualité à différents niveaux de détails grâce à des fonctions OLAP. Il est alors possible de visualiser l'information de qualité, de l'indicateur global aux données sources servant au calcul des indicateurs, sans avoir de surcharge d'information.

5. Conclusion et perspectives

Cet article présente une nouvelle approche permettant de communiquer l'information relative à la qualité des données géospatiales dans le but de réduire les risques de mauvaises utilisations. Afin de ne pas surcharger les utilisateurs d'informations et de supporter adéquatement leur processus de décision, cette approche préconise l'utilisation de tableaux de bord et d'indicateurs de qualité intégrés dans l'interface du SIG. La qualité étant ici définie comme « l'adéquation à l'utilisation », l'information fournie aux utilisateurs compare les attentes de ceux-ci aux spécifications et caractéristiques internes des données. Cette information relative à la qualité des données peut être basée sur les métadonnées ou toute autre source d'information sur la qualité. L'information sur la qualité est alors communiquée à l'utilisateur sous la forme d'indicateurs de statut ou de risque que celui-ci peut sélectionner, modifier au besoin, puis consulter à différents niveaux de détails. Étant donné l'hétérogénéité spatiale de l'information sur la qualité, des outils permettant une visualisation cartographique de la qualité sont également proposés. Cette approche fournie aux utilisateurs de SIG des outils leur permettant d'identifier rapidement des divergences potentielles entre leurs besoins tels qu'exprimés et la qualité des données telle que documentée. Une telle approche peut être intégrée dans des outils SIG ou dans d'autres outils de visualisation cartographique (ex : SOLAP), soit comme un outil de gestion de la qualité à part entière, soit comme une composante de l'outil de visualisation parmi d'autres, pouvant être activée par l'usager au besoin.

Les métadonnées définies par les organismes de normalisation et actuellement fournies par les producteurs de données sont nécessaires pour permettre la création des indicateurs, mais présentent des limites dans leur format actuel. En effet, de nombreuses métadonnées utilisent des textes libres pour décrire les données, ce type de format étant difficilement manipulable automatiquement. De plus, les métadonnées offrent la plupart du temps des descriptions au niveau du jeu de données uniquement. Pour tirer le maximum de bénéfices du système MUM les données devraient décrire les objets à un niveau de détail plus fin afin de mieux souligner l'hétérogénéité spatiale, temporelle ou descriptive de la qualité. Beaucoup de jeux de données n'ont pas de métadonnées ou ont des métadonnées sommaires. Toutefois, pas d'information est en soi une information utile à l'utilisateur des données, lui indiquant que les données sont peu documentées et que leur utilisation peut donc être délicate. L'utilisateur peut alors décider de réduire son incertitude en acquérant des informations complémentaires sur les jeux de données ou de travailler avec ces données en absorbant ainsi l'incertitude résiduelle.

Remerciements

Ce travail est financé par le Ministère de la Recherche, de la Science et de la Technologie du Québec dans le cadre de la collaboration avec le projet européen REVIGIS, le Centre de Recherche en Géomatique (CRG) et l'Université Laval. Nous remercions également le Centre d'Information Topographique de Sherbrooke (CIT-S) de Géomatique Canada pour leur support ainsi que des évaluateurs anonymes pour leurs commentaires.

6. Bibliographie

Agumya A., Hunter G. J., « Determining fitness for use of geographic information », *ITC Journal*, vol. 2, nº 1, 1997, p. 109-113.

Beard K., « Use error: the neglected error component », *Proceedings of AUTO-CARTO 9*, Baltimore, Maryland, mars 1989, p. 808-817.

Bédard Y., « Uncertains in Land Information Systems Databases », *Proceedings of Eighth International Symposium on Computer-Assisted Cartography*, Baltimore, Maryland, 29 mars - 3 avril 1986, American Society for Photogrammetry and Remote Sensing et American Congress on Surveying and Mapping, p. 175-184.

Blackmore M., « High or Low Resolution? Conflicts of Accuracy, Cost, Quality and Application in Computer Mapping », *Computers & Geosciences*, vol. 11, nº 2, 1985, p. 345-348.

Buttenfield B. P., « Representing Data Quality », *Cartographica*, vol. 30, nº 2-3, 1993, p. 1-7.

Curry M. R., *Digital places: Living with Geographic Information Technologies*, London & New-York, Routeledge, 1998.

Devillers R. Bédard Y., Jeansoulin R., « Multidimensional management of geospatial data quality information for its dynamic use within Geographical Information Systems », *Photogrammetric Engineering & Remote Sensing (PE&RS)*, 2004, soumis.

Duckham M., McCreadie J., « An intelligent, distributed, error-aware OOGIS », *Proceedings of 1st International Symposium on Spatial Data Quality*, Hong Kong, 18-20 juillet 1999, p. 496-506.

Duckham M., McCreadie J. E., « Error-aware GIS Development ». *Spatial Data Quality* (W. Shi, P. F. Fisher et M. F. Goodchild, Eds), Taylor & Francis, London, 2002, p. 63-75.

Elshaw Thrall S., Thrall G. I., « Desktop GIS software ». *Geographical Information Systems* (P. A. Longley, M. F. Goodchild, D. J. Maguire et D. W. Rhind, Eds), John Wiley & Sons, New-York, 1999, p. 331-345.

Epstein E. F. Hunter G. J., Agumya A., « Liability insurance and the use of geographical information », *International Journal of Geographical Information Science*, vol. 12, nº 3, 1998, p. 203-214.

Faïz S. O., *Systèmes d'Informations Géographiques: Information Qualité et Data Mining*, Tunis, Éditions C.L.E., 1999.

Fernandez A., *Les nouveaux tableaux de bord des décideurs*, Paris, Éditions d'organisation, 2000.

FGDC, Content Standard for Digital Geospatial Metadata Workbook version 2, 2000.

Fisher G., « User Modeling in Human-Computer Interaction », *User Modeling and User-Adapted Interaction*, vol. 11, nº, 2001, p. 65-86.

Frank A. U., « Metamodels for Data Quality Description ». *Data Quality in Geographic Information - From Error to Uncertainty* (M. F. Goodchild et R. Jeansoulin, Eds), Editions Hermès, 1998, p. 192.

Fisher P., « Models of uncertainty in spatial data ». *Geographical Information Systems* (P. A. Longley, M. F. Goodchild, D. J. Maguire et D. W. Rhind, Eds), John Wiley & Sons, Inc., 1999, p. 191-205.

Gan E., Shi W., « Error Metadata Management System ». *Spatial Data Quality* (W. Shi, P. F. Fisher et M. F. Goodchild, Eds), Taylor Francis, London and New York, 2002, p. 336.

Gervais M., *Pertinence d'un manuel d'instructions au sein d'une stratégie de gestion du risque juridique découlant de la fourniture de données géographiques numériques*, Thèse de doctorat, Université Laval, Québec, 2004.

Goodchild M. F., « Sharing Imperfect Data ». *Sharing Geographic Information* (H. J. Onsrud et G. Rushton, Eds), Rutgers University Press, New Brunswick, NJ, p. 413-425, 1995.

Goodchild M. F., Kemp K. K., NCGIA Core Curriculum in GIS, National Center for Geographic Information and Analysis, University of California, Santa Barbara CA, 1990.

Guptill S. C., Morrison J. L., *Elements of spatial data quality*, New York, Elsevier Science, 1995.

Harvey F., « Quality Needs More Than Standards ». *Data Quality in Geographic Information - From Error to Uncertainty* (M. F. Goodchild et R. Jeansoulin, Eds), Editions Hermès, 1998, p. 192.

Hunter G. J., « Managing uncertainty in GIS ». *Geographical Information Systems* (P. A. Longley, M. F. Goodchild, D. J. Maguire et D. W. Rhind, Eds), John Wiley & Sons, Inc., 1999, p. 633-641.

Hunter G. J., « Spatial Data Quality Revisited ». *Proceedings of GeoInfo 2001 Symposium*, Rio de Janeiro, Brésil, 4-5 octobre 2001.

Hunter G. J., Reinke K. J., « Adapting Spatial Databases to Reduce Information Misuse Through Illogical Operations », *Proceedings of 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences (Accuracy 2000)*, Amsterdam, juillet 2000, p. 313-319.

ISO-TC/211, Geographic Information - Quality principles 19113, 2003.

Kaplan R., Norton D., « The balanced scorecard: Measures that Drive Performance », *Harvard Business Review*, vol. 70, nº 1, 1992, p. 71-79.

Klein G., *Sources of Power - How people make decisions*, Cambridge, Massachusetts, MIT Press, 1999.

Krek A., Frank A. U., « Optimization of Quality of Geoinformation Products », *Proceedings of Proceedings of 11th Annual Colloquium of the Spatial Information Research Centre, SIRC'99*, Dunedin, New Zealand, 13-15 décembre, 1999, Department of Information Science, University of Otago, p. 151-159.

Lardon S., Maurel P., Piveteau V., ed., 2001. *Représentations spatiales et développement territorial*, Éditions Hermès.

Longley P. A., Goodchild M. F., Maguire D. J., Rhind D. W., ed., 1999. *Geographical Information Systems*, John Wiley & Sons.

Martinet B., Marti Y.-M., *L'intelligence économique*, Éditions d'Organisation, 2001.

Miller G. A., « The Magical Number Seven, plus or minus two: Some limits on our capacity for processing information », *The Psychological Review*, vol. 63, nº, 1956, p. 81-97.

Miller H. J., Han J., *Geographic Data mining and Knowledge Discovery*, Taylor & Francis, 2001.

Mintzberg H., *The structuring of organisations*, Englewood Cliffs, Prentice-Hall, 1979.

Monmonier M., « A Case Study in the Misuse of GIS: Siting a Low-Level Radioactive Waste Disposal Facility in New-York Sate », *Proceedings of Conference on Law and Information Policy for Spatial Databases*, Tempe (USA), 1994, p. 293-303.

Plan Canada, Sustainable community indicators programme, vol 39, nº 5, 1999.

Qiu J., Hunter G. J., « A GIS with the Capacity for Managing Data Quality Information ». *Spatial Data Quality* (W. Shi, M. F. Goodchild et P. F. Fisher, Eds), Taylor & Francis, London, 2002, p. 230-250.

Reinke K. J., Hunter G. J., « A Theory for Communicating Uncertainty in Spatial Databases ». *Spatial Data Quality* (W. Shi, P. F. Fisher et M. F. Goodchild, Eds), Taylor & Francis, London, 2002, p. 77-101.

Rivest S., Bédard Y., Marchand P., « Towards Better Support for Spatial Decision Making: Defining the Characteristics of Spatial On-Line Analytical Processing (SOLAP) », *Geomatica*, vol. 55, nº 4, 2001, p. 539-555.

Shannon C. E., « A Mathematical Theory of Communication », *The Bell System Technical Journal*, vol. 27, nº, 1948, p. 379-423.

Timpf S., Raubal M., Kuhn, W., « Experiences with Metadata », *Proceedings of Symposium on Spatial Data Handling, SDH'96, Advances in GIS Research II*, Delft, The Netherlands, 12-16 août 1996, IGU, p. 12B.31 - 12B.43.

Voyer P., *Tableaux de bord de gestion et indicateurs de performance*, Presse de l'Université du Québec, 2000.